Learning Correlation-aware Aleatoric Uncertainty for 3D Hand Pose Estimation

Lee Chae-Yeon[1] Nam Hyeon-Woo[1] Tae-Hyun Oh[2]
[1]POSTECH [2]KAIST

BMVC 2025

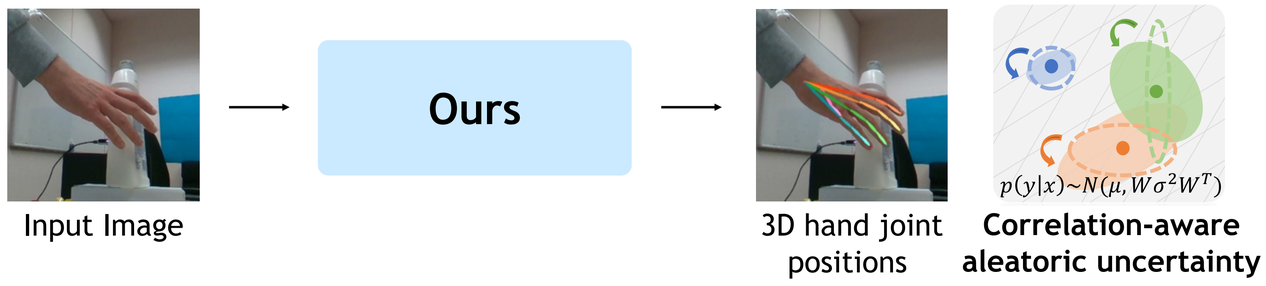
1. Summary

Limitation in 3D Hand Pose Estimation

- Inability to model aleatoric (data) uncertainty
- Uncertainty modeling in the absence of joint correlation knowledge

Contributions

- Introduce aleatoric uncertainty modeling into 3D hand pose estimation
- Achieve a better trade-off btw. modeling joint correlations & efficiency



2. Preliminary

A. No uncertainty modeling: Deterministic

$$g(f(\mathbf{x})) = \boldsymbol{\mu}$$

- g : $\mathbb{R}^{d_f} \rightarrow \mathbb{R}^{d_o}$

B. Uncertainty modeling w/ diagonal cov. matrix

parameter-efficient ✗ correlation modeling

$$g_{diag}(f(\mathbf{x})) \sim \mathcal{N}(\boldsymbol{\mu}, diag(\boldsymbol{\sigma}^2))$$

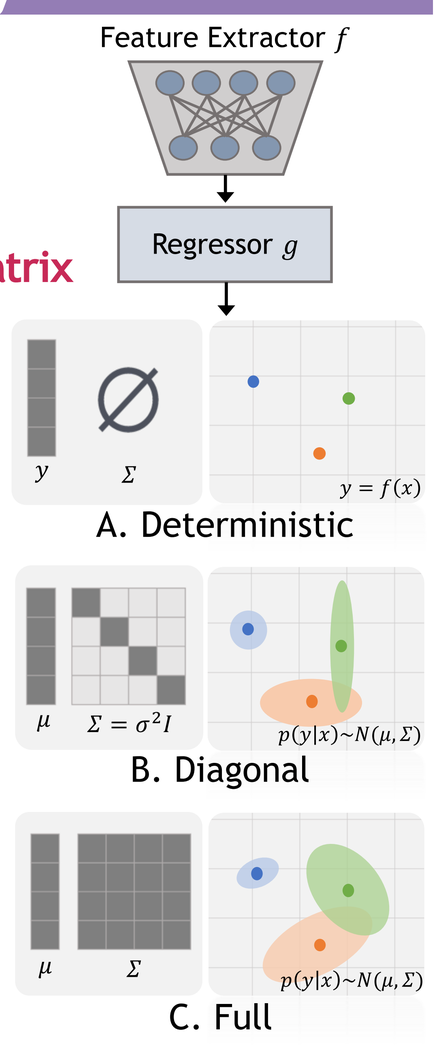
- g : $\mathbb{R}^{d_f} \rightarrow \mathbb{R}^{2 \cdot d_o}$
- $\mathcal{L}_{NLL} = \log p(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\sigma}^2)) = \frac{\|\boldsymbol{y}-\boldsymbol{\mu}\|^2}{2\boldsymbol{\sigma}^2} + 0.5 \log \boldsymbol{\sigma}^2$

C. Uncertainty modeling w/ full cov. matrix

✗ parameter-efficient ○ correlation modeling

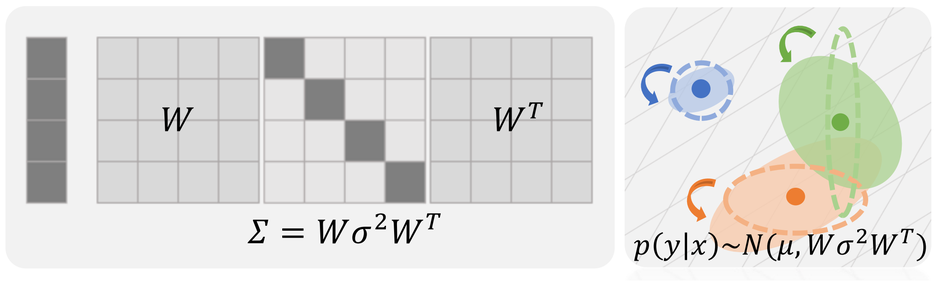
$$g_{full}(f(\mathbf{x})) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

- g : $\mathbb{R}^{d_f} \rightarrow \mathbb{R}^{d_o + d_o^2}$
- $\mathcal{L}_{NLL} = \log p(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma})) = \frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}-\boldsymbol{\mu}) + 0.5 \log\det \boldsymbol{\Sigma}$



3. Method

Key Idea: Mid-representation between Diagonal and Full



Independence-assumed space → correlation-aware space
using **a single linear layer of weights W**

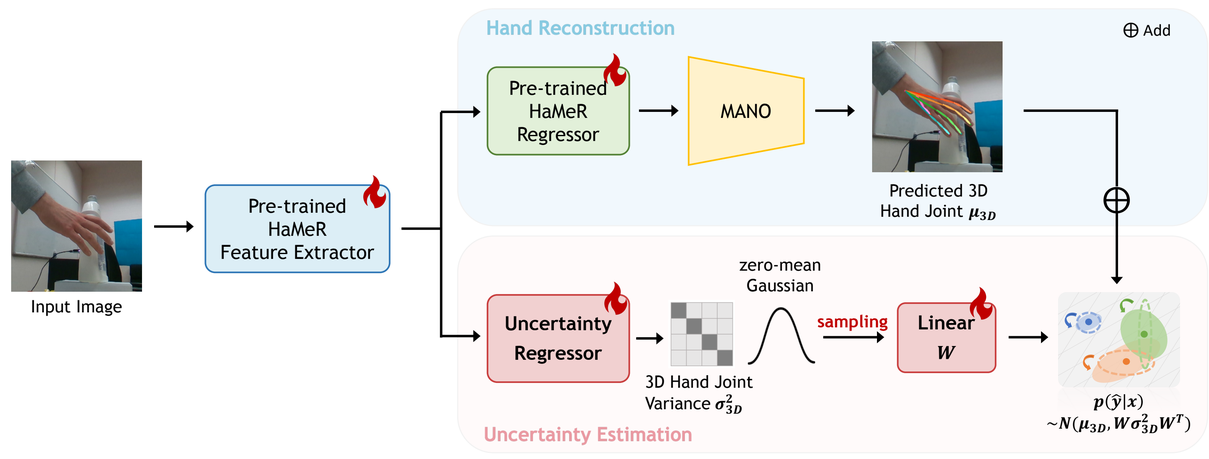
Parametrization	Uncertainty modeling	# Params.	Example [# Params]
Deterministic	✗	$d_f d_o$	0.065M
Diagonal covariance	✓(independent)	$2d_f d_o$	0.129M
Full covariance	✓	$d_f(d_o + d_o^2)$	4.129M
Ours	✓	$2d_f d_o + d_o^2$	0.133M

→ # of parameters: diagonal ≈ ours ≪ full

d_f = 1024, d_o = 63

→ Expressiveness in capturing correlations: diagonal < ours < full

Overall Pipeline: Adopting HaMeR



1. Add uncertainty regressor to output per-joint variances
2. **Independence-assumed space**: formulate zero-mean Gaussian dist.
3. Transform into **correlation-aware space**: Draw N samples → apply a single linear transformation → add mean

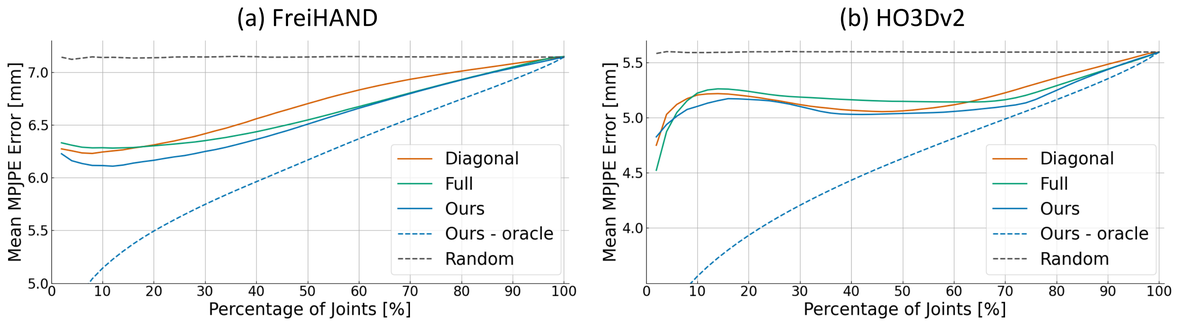
$$\mathrm{p}(\mathbf{z}|\mathbf{x}) \sim \mathcal{N}\left(\mathbf{0}, diag(\boldsymbol{\sigma}^2_{\boldsymbol{3D}})\right) \rightarrow \mathrm{p}(\hat{\mathbf{y}}|\mathbf{x}) \sim \mathcal{N}\left(\mathbf{0}, \mathbf{W} diag(\boldsymbol{\sigma}^2_{\boldsymbol{3D}}) \mathbf{W}^{\mathrm{T}}\right)$$
$$\rightarrow \mathrm{p}(\hat{\mathbf{y}}|\mathbf{x}) \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{3D}}, \mathbf{W} diag(\boldsymbol{\sigma}^2_{\boldsymbol{3D}}) \mathbf{W}^{\mathrm{T}}\right)$$

[1] Pavlakos et al., "Reconstructing Hands in 3D with Transformers," CVPR 2024.

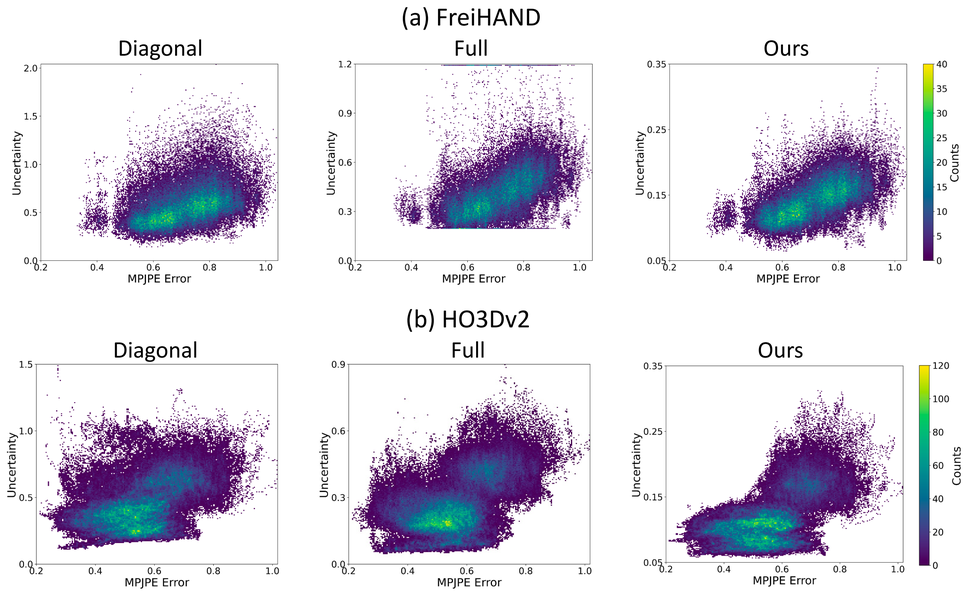
4. Experiments

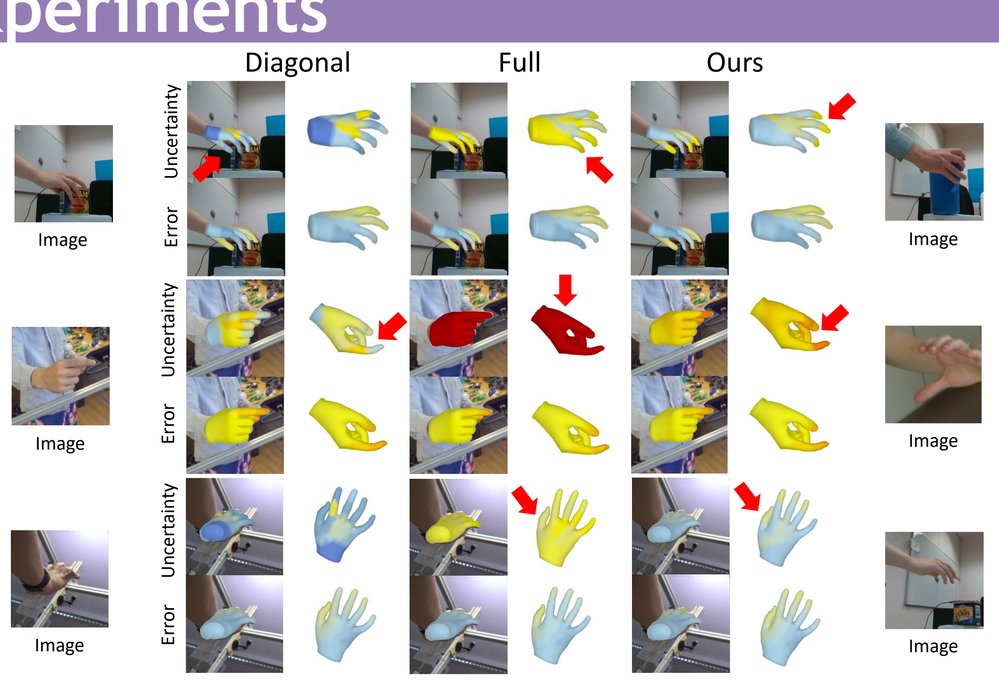
Qualitative Results

- **Sparsification curve**: sorted based on estimated uncertainty and top x% most certain joints are evaluated
  - Oracle (GT): sorted by actual MPJPE error, Random: sorted randomly



- **2D histogram** btw. MPJPE error & estimated uncertainty





Quantitative Results

- (Top) Achieve superior uncertainty estimation performance
- (Bottom) Maintain competitive 3D hand pose estimation performance

Method	FreiHAND			HO3Dv2		
	AUSC ↓	AUSE ↓	Pearson's ρ ↑	AUSC ↓	AUSE ↓	Pearson's ρ ↑
Diagonal	655	54.6	0.393	511	63.3	0.448
Full	648	47.6	0.453	512	64.3	0.493
Ours	**642**	**42.2**	**0.569**	**505**	**57.6**	**0.600**

Method	FreiHAND				HO3Dv2					
	PA-MPJPE ↓	PA-MPVPE ↓	F@5 ↑	F@15 ↑	AUC_J ↑	PA-MPJPE ↓	AUC_V ↑	PA-MPVPE ↓	F@5 ↑	F@15 ↑
HaMeR	6.0	**5.7**	**0.785**	**0.990**	0.846	7.7	0.841	7.9	0.635	0.980
Deterministic	6.1	**5.7**	**0.785**	**0.990**	0.845	7.8	0.840	8.0	0.629	0.980
Diagonal	6.1	5.8	0.782	**0.990**	0.847	7.7	0.842	7.9	0.638	**0.981**
Full	6.1	5.9	0.773	0.989	**0.849**	**7.6**	**0.844**	**7.8**	**0.644**	**0.981**
Ours	6.0	**5.7**	0.784	**0.990**	0.847	**7.6**	0.843	7.9	0.640	**0.981**